

B4

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

| | | | |
|--|--------------------------|--|--|
| (51) International Patent Classification ⁷ : C12N 15/12, C07K 14/47, 16/18, G01N 33/566, C12Q 1/68, C12N 15/11, 15/62, A01K 67/027, A61K 38/00 | | A2 | (11) International Publication Number: WO 00/58473 |
| | | | (43) International Publication Date: 5 October 2000 (05.10.00) |
| (21) International Application Number: PCT/US00/08621 | | (72) Inventors; and | |
| (22) International Filing Date: 31 March 2000 (31.03.00) | | (75) Inventors/Applicants (for US only): SHIMKETS, Richard, A. [US/US]; 191 Leete Street, West Haven, CT 06516 (US). LEACH, Martin [GB/US]; 884 School Street, Webster, MA 01570 (US). | |
| (30) Priority Data: | | (74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Cohn, Ferris, Glovsky and Popeo, P.C., One Financial Center, Boston, MA 02111 (US). | |
| 60/127,607 | 31 March 1999 (31.03.99) | US | (81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). |
| 60/127,636 | 2 April 1999 (02.04.99) | US | |
| 60/127,728 | 5 April 1999 (05.04.99) | US | |
| 09/540,763 | 30 March 2000 (30.03.00) | US | |
| (63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications | | | |
| US | 60/127,607 (CIP) | | Published Without international search report and to be republished upon receipt of that report. |
| Filed on | 31 March 1999 (31.03.99) | | |
| US | 60/127,636 (CIP) | | |
| Filed on | 2 April 1999 (02.04.99) | | |
| US | 60/127,728 (CIP) | | |
| Filed on | 5 April 1999 (05.04.99) | | |
| US | 09/540,763 (CIP) | | |
| Filed on | 30 March 2000 (30.03.00) | | |
| (71) Applicant (for all designated States except US): CURAGEN CORPORATION [US/US]; 555 Long Wharf Drive, 11th Floor, New Haven, CT 06511 (US). | | | |

(54) Title: NUCLEIC ACIDS INCLUDING OPEN READING FRAMES ENCODING POLYPEPTIDES; "ORFX"

(57) Abstract

The present invention provides open reading frames ORFX, encoding isolated polypeptides, as well as polynucleotides encoding ORFX and antibodies that immunospecifically bind to ORFX or any derivative, variant, mutant, or fragment of the ORFX polypeptides, polynucleotides or antibodies. The invention additionally provides methods in which the ORFX polypeptide, polynucleotide and antibody are used in detection and treatment of a broad range of pathological states, as well as to other uses.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | | | |
|----|--------------------------|----|--|----|--|----|--------------------------|
| AL | Albania | ES | Spain | LS | Lesotho | SI | Slovenia |
| AM | Armenia | FI | Finland | LT | Lithuania | SK | Slovakia |
| AT | Austria | FR | France | LU | Luxembourg | SN | Senegal |
| AU | Australia | GA | Gabon | LV | Latvia | SZ | Swaziland |
| AZ | Azerbaijan | GB | United Kingdom | MC | Monaco | TD | Chad |
| BA | Bosnia and Herzegovina | GE | Georgia | MD | Republic of Moldova | TG | Togo |
| BB | Barbados | GH | Ghana | MG | Madagascar | TJ | Tajikistan |
| BE | Belgium | GN | Guinea | MK | The former Yugoslav Republic of Macedonia | TM | Turkmenistan |
| BF | Burkina Faso | GR | Greece | | | TR | Turkey |
| BG | Bulgaria | HU | Hungary | ML | Mali | TT | Trinidad and Tobago |
| BJ | Benin | IE | Ireland | MN | Mongolia | UA | Ukraine |
| BR | Brazil | IL | Israel | MR | Mauritania | UG | Uganda |
| BY | Belarus | IS | Iceland | MW | Malawi | US | United States of America |
| CA | Canada | IT | Italy | MX | Mexico | UZ | Uzbekistan |
| CF | Central African Republic | JP | Japan | NE | Niger | VN | Viet Nam |
| CG | Congo | KE | Kenya | NL | Netherlands | YU | Yugoslavia |
| CH | Switzerland | KG | Kyrgyzstan | NO | Norway | ZW | Zimbabwe |
| CI | Côte d'Ivoire | KP | Democratic People's Republic of Korea | NZ | New Zealand | | |
| CM | Cameroon | | | PL | Poland | | |
| CN | China | KR | Republic of Korea | PT | Portugal | | |
| CU | Cuba | KZ | Kazakstan | RO | Romania | | |
| CZ | Czech Republic | LC | Saint Lucia | RU | Russian Federation | | |
| DE | Germany | LI | Liechtenstein | SD | Sudan | | |
| DK | Denmark | LK | Sri Lanka | SE | Sweden | | |
| EE | Estonia | LR | Liberia | SG | Singapore | | |

What is claimed is:

1. An isolated nucleic acid molecule encoding a polypeptide comprising an amino acid sequence that is at least 85% identical to a polypeptide including an amino acid sequence selected from the group consisting of SEQ ID NO:2 n , wherein n is any integer 1-3161, or the complement thereof.
2. The isolated nucleic acid molecule of claim 1, said molecule hybridizing under stringent conditions to a nucleic acid sequence complementary to a nucleic acid molecule comprising the sequence of nucleotides selected from the group consisting of SEQ ID NO:2 n , wherein n is any integer 1-3161, or the complement thereof.
3. The isolated nucleic acid molecule of claim 1, said molecule encoding a polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO: 2 n , wherein n is any integer 1-3161, or an amino acid sequence comprising one or more conservative substitutions in the amino acid sequence selected from the group consisting of SEQ ID NO: 2 n .
4. The isolated nucleic acid molecule of claim 1, wherein said molecule encodes a polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO: 2 n , wherein n is any integer 1-3161.
5. The isolated nucleic acid molecule of claim 1, wherein said molecule comprises the sequence of nucleotides selected from the group consisting of SEQ ID NO:2 n -1, wherein n is any integer 1-3161, or the complement thereof.
6. An oligonucleotide less than 100 nucleotides in length and comprising at least contiguous nucleotides selected from the group consisting of SEQ ID NO:2 n -1, wherein n is any integer 1-3161, or the complement thereof.
7. A vector comprising the nucleic acid molecule of claim 1.

8. The vector of claim 7, wherein said vector is an expression vector.
9. A host cell comprising the isolated nucleic acid molecule of claim 1.
10. A substantially purified polypeptide comprising an amino acid sequence at least 80% identical to a polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO: 2*n*, wherein *n* is any integer 1-3161.
11. The polypeptide of claim 10, wherein said polypeptide comprises the amino acid sequence selected from the group consisting of SEQ ID NO: 2*n*, wherein *n* is any integer 1-3161.
12. An antibody that selectively binds to the polypeptide of claim 10.
13. A pharmaceutical composition comprising a therapeutically or prophylactically effective amount of a therapeutic selected from the group consisting of:
 - a) the nucleic acid of claim 1;
 - b) the polypeptide of claim 10; and
 - c) the antibody of claim 12;and a pharmaceutically acceptable carrier.
14. A kit comprising in one or more containers, a therapeutically or prophylactically effective amount of the pharmaceutical composition of claim 13.
15. A method of producing the polypeptide of claim 10, said method comprising culturing the host cell of claim 9 under conditions in which the nucleic acid molecule is expressed.
16. A method of detecting the presence of the polypeptide of claim 10 in a sample, comprising contacting the sample with a compound that selectively binds to said polypeptide under conditions allowing the formation of a complex between said polypeptide and said

compound, and detecting said complex, if present, thereby identifying said polypeptide in said sample.

17. A method of detecting the presence of a nucleic acid molecule of claim 1 in a sample, the method comprising contacting the sample with a nucleic acid probe or primer that selectively binds to the nucleic acid molecule and determining whether the nucleic acid probe or primer bound to the nucleic acid molecule of claim 1 is present in the sample.

18. A method for modulating the activity of the polypeptide of claim 10, the method comprising contacting a cell sample comprising the polypeptide of claim 10 with a compound that binds to said polypeptide in an amount sufficient to modulate the activity of the polypeptide.

19. The use of a therapeutic in the manufacture of a medicament for treating a syndrome associated with a ORFX-associated disorder, wherein said therapeutic is selected from the group consisting of:

- a) the nucleic acid of claim 1;
- b) the polypeptide of claim 10; and
- c) the antibody of claim 12.

20. A method for screening for a modulator of activity or of latency or predisposition to an ORFX-associated disorder, said method comprising:

- a) contacting a test compound with the polypeptide of claim 10; and
- b) determining if said test compound binds to said polypeptide,

wherein binding of said test compound to said polypeptide indicates the test compound is a modulator of activity or of latency or predisposition to an ORFX-associated disorder.

21. A method for screening for a modulator of activity or of latency or predisposition to an ORFX-associated disorder, said method comprising:

- a) administering a test compound to a test subject at an increased risk ORFX-associated disorder, wherein said test subject recombinantly expresses a polypeptide encoded by the nucleotide of claim 1;

- b) measuring expression the activity of said protein in said test subject;
- c) measuring the activity of said protein in a control subject that recombinantly expresses said protein and is not at increased risk for an ORFX-associated disorder; and
- d) comparing expression of said protein in said test subject and said control subject, wherein a change in the activity of said protein in said test subject relative to said control subject indicates the test compound is a modulator or of latency of predisposition to an ORFX-associated disorder.

22. The method of claim 20, wherein said test animal is a recombinant test animal that expresses a test protein transgene or expresses said transgene under the control of a promoter at an increased level relative to a wild-type test animal, and wherein said promoter is not the native gene promoter of said transgene.

23. A method for determining the presence of or predisposition to a disease associated with altered levels of a polypeptide of claim 11 in a subject, the method comprising:

- a) measuring the amount of the polypeptide in a sample from said subject; and
- b) comparing the amount of said polypeptide in step (a) to the amount of the polypeptide present in a control sample,

wherein an alteration in the level of the polypeptide in step (a) as compared to the control sample indicates the presence of or predisposition to a disease in said subject.

24. The method of claim 23, wherein said subject is a human.

25. A method for determining the presence of or predisposition to a disease associated with altered levels the nucleic acid molecule of claim 1 in a subject, the method comprising:

- a) measuring the amount of the nucleic acid in a sample from the mammalian subject; and
- b) comparing the amount of said nucleic acid in step (a) to the amount of the nucleic acid present in a control sample,

wherein an alteration in the level of the nucleic acid in step (a) as compared to the control sample indicates the presence of or predisposition to said disease in said subject.

26. The method of claim 25, wherein said subject is a human.

27. A method of treating or preventing a pathological condition associated with an ORFX-associated disorder in a subject, the method comprising administering to said subject a polypeptide of claim 10 in an amount sufficient to alleviate or prevent said pathological condition.

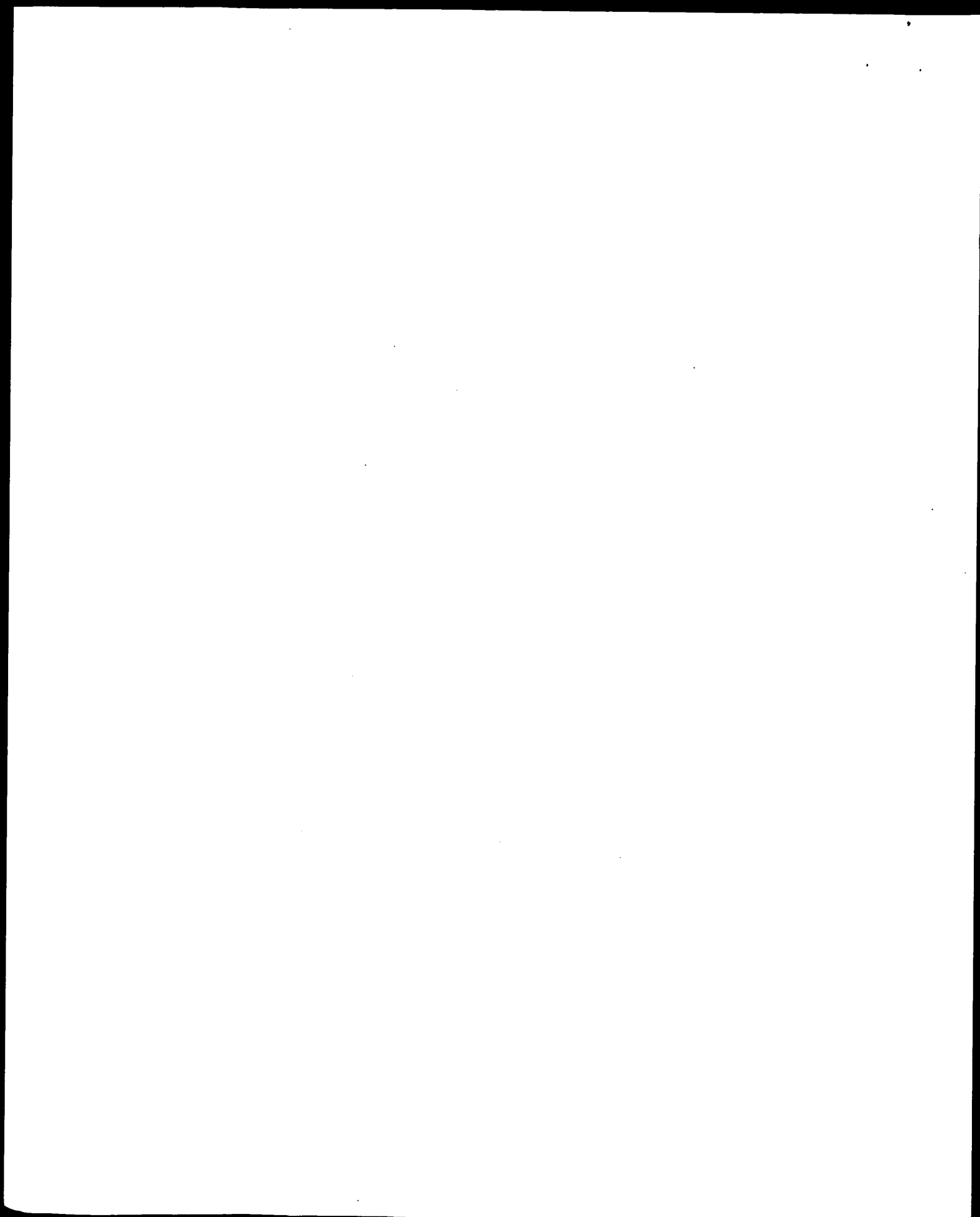
28. The method of claim 27, wherein said subject is a human.

29. A method of treating or preventing a pathological condition associated with an ORFX-associated disorder in a subject, the method comprising administering to said subject a nucleic acid molecule of claim 1 in an amount sufficient to alleviate or prevent said pathological condition.

30. The method of claim 29, wherein said subject is a human.

31. A method of treating or preventing a pathological condition associated with an ORFX-associated disorder in a subject, the method comprising administering to said subject an antibody of claim 12 in an amount sufficient to alleviate or prevent said pathological condition.

32. The method of claim 31, wherein said subject is a human.



<213> Homo sapiens

<400> 360

```

Thr Arg Thr Asp Lys Arg Pro Val Met Ala Asp Leu Arg Glu Ser Gly
 1           5           10           15
Ala Ile Glu Gln Asp Ala Asp Met Ile Val Phe Ile Tyr Arg Asp Asp
          20           25           30
Tyr Tyr Asn Lys Glu Asn Ser Pro Asp Lys Gly Leu Ala Glu Ile Ile
          35           40           45
Ile Gly Lys His Arg Gly Gly Pro Thr Gly Ser Cys Lys Leu Lys Phe
          50           55           60
Phe Gly Glu Tyr Thr Arg Phe Asp Asn Leu Ala His Asn Ser Val Gly
65           70           75           80
Ser Phe Glu

```

<210> 361

<211> 453

<212> DNA

<213> Homo sapiens

<400> 361

```

gctttgcagg aggaaatctc tatctctggc tgcaagatga ggctgagcta cctgagcagc
60
cggacccttg gctacaaatc tgtcctgagg atcagcctca cccaccgac catcccttc
120
aacctcatga aggtgcacct catggtagcg gtggagggcc gcctcttcag gaagtgggtc
180
gctgcagccc cagacctgtc ctattatttc atttgggaca agacagacgt ctacaaccag
240
aaggtgtttg ggctttcaga agcctttggt tccgtgggtt atgaatatga atcctgcccc
300
gatctaattc tgtgggaaaa aagaacaaca gtgctgcagg gctatgaaat tgacgcgtcc
360
aagcttggag gatggagcct agacaaacat catgccctca acattcaaag tggcatcctg
420
cacaaagga atgngagaa ccagtttgtg tct
453

```

<210> 362

<211> 151

<212> PRT

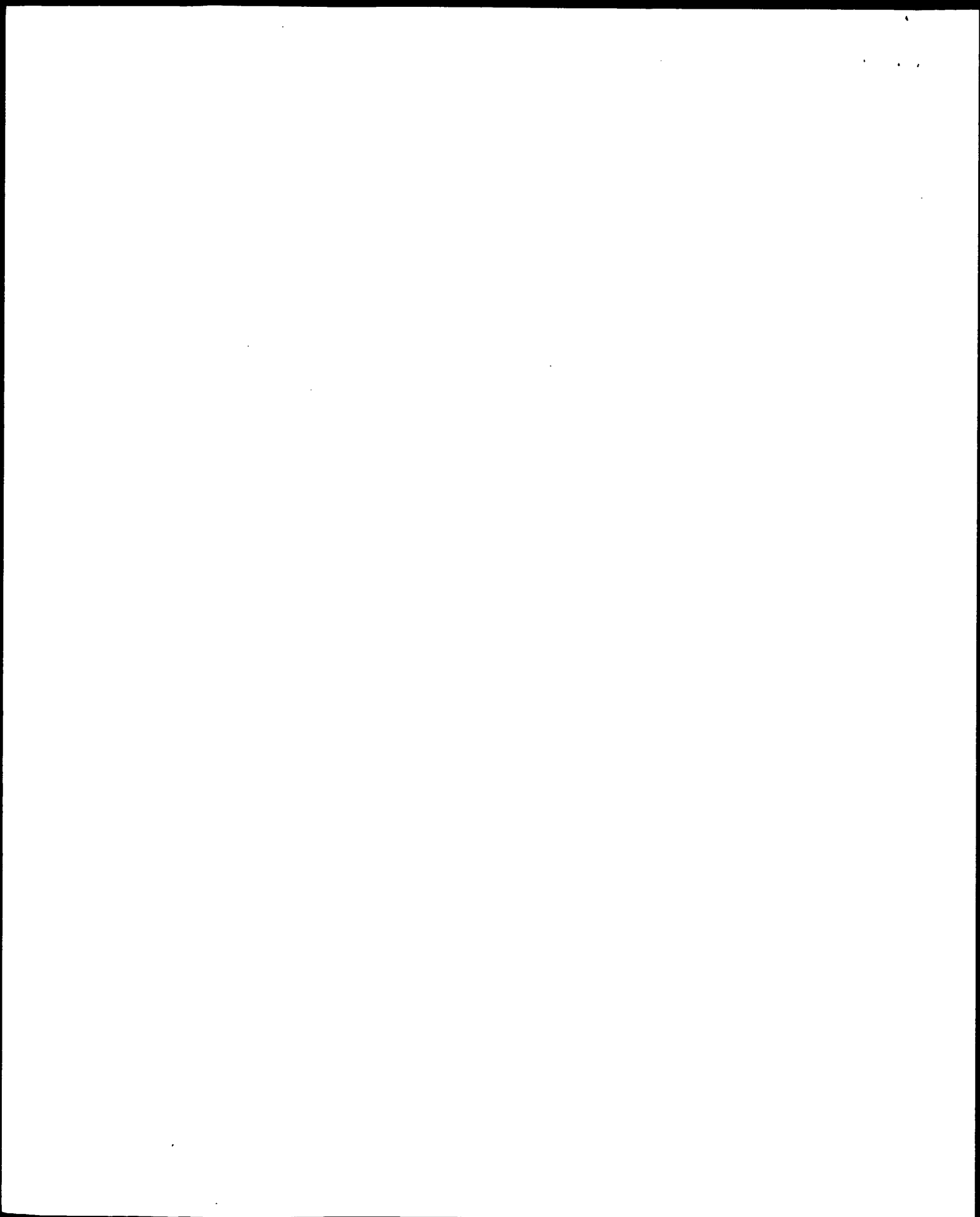
<213> Homo sapiens

<400> 362

```

Ala Leu Gln Glu Glu Ile Ser Ile Ser Gly Cys Lys Met Arg Leu Ser
 1           5           10           15
Tyr Leu Ser Ser Arg Thr Pro Gly Tyr Lys Ser Val Leu Arg Ile Ser
          20           25           30
Leu Thr His Pro Thr Ile Pro Phe Asn Leu Met Lys Val His Leu Met
          35           40           45
Val Ala Val Glu Gly Arg Leu Phe Arg Lys Trp Phe Ala Ala Ala Pro
          50           55           60
Asp Leu Ser Tyr Tyr Phe Ile Trp Asp Lys Thr Asp Val Tyr Asn Gln

```



65 70 75 80
 Lys Val Phe Gly Leu Ser Glu Ala Phe Val Ser Val Gly Tyr Glu Tyr
 85 90 95
 Glu Ser Cys Pro Asp Leu Ile Leu Trp Glu Lys Arg Thr Thr Val Leu
 100 105 110
 Gln Gly Tyr Glu Ile Asp Ala Ser Lys Leu Gly Gly Trp Ser Leu Asp
 115 120 125
 Lys His His Ala Leu Asn Ile Gln Ser Gly Ile Leu His Lys Gly Asn
 130 135 140
 Gly Glu Asn Gln Phe Val Ser
 145 150

<210> 363
 <211> 502
 <212> DNA
 <213> Homo sapiens

<400> 363
 ggtaccaaaa aagtttgcca cagtattcac actccagggtc tccataaacc ttccagatcc
 60
 gctcacacaa gctgggtgttc atttgcttct tctgtaaact gttcaggacc ttcattgaaag
 120
 cggtagtgcc tgaccggtgc tcaggggagc ctttgcaaga gtcaggctga tgtgtgatgg
 180
 tgtccccacc accagctact ggagggagga ggtctgaggg ctcagctggg tttgacctga
 240
 gacacctgct gggatctggg tcaccagctg aaagcacagc catgttctgc cttcccccta
 300
 gggggctctg ggcgccatgg ctttctgat ctgaccagc actctgggccc ttggacagca
 360
 gtagtgtgat cacttcacct tgcgtctgga ctgagcttct gtgctgcatg tctgggggct
 420
 tctcaggagc agcatgagcc tctgcggagg aggtatcatt tttcaacaaa aaatcatctg
 480
 aaaccacctc ttgagaatgc ag
 502

<210> 364
 <211> 136
 <212> PRT
 <213> Homo sapiens

<400> 364
 Met Gln His Arg Ser Ser Val Gln Thr Gln Gly Glu Val Ile Thr Leu
 1 5 10 15
 Leu Leu Ser Lys Ala Gln Ser Ala Gly Ser Asp Gln Glu Ser His Gly
 20 25 30
 Ala Gln Ser Pro Leu Gly Glu Gly Gln Asn Met Ala Val Leu Ser Ala
 35 40 45
 Gly Asp Pro Asp Pro Ser Arg Cys Leu Arg Ser Asn Pro Ala Glu Ala
 50 55 60
 Ser Asp Leu Leu Pro Pro Val Ala Gly Gly Gly Asp Thr Ile Thr His
 65 70 75 80
 Gln Pro Asp Ser Cys Lys Ala Ala Pro Glu His Arg Ser Gly Ile Thr

ID HSM010618 standard; RNA; EST; 718 BP.

XX
AC AL045768;

XP-002237532

XX
SV AL045768.1

XX
DT 12-MAR-1999 (Rel. 59, Created)
DT 20-FEB-2000 (Rel. 62, Last updated, Version 4)

P.D. 20-02-2000

P. 1

XX
DE Homo sapiens mRNA; EST DKFZp434F206_r1 (from clone DKFZp434F206)

XX
KW EST; expressed sequence tag.

XX
OS Homo sapiens (human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

XX
RN [1]

RP 1-718

RA Koehrer K., Beyer A., Mewes H.W., Gassenhuber J., Wiemann S.;

RT ;

RL Submitted (18-FEB-2000) to the EMBL/GenBank/DDBJ databases.

RL MIPS, Am Klopferspitz 18a, D-82152 Martinsried, GERMANY

XX

DR RZPD; DKFZp434F206; DKFZp434F206.

DR UNILIB; 1752; 1752.

XX

CC This is the 5' sequence of the clone insert

CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

CC Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

CC sequenced by BMFZ (Biomedical Research Center at the Charite,

CC Berlin/Germany) within the cDNA sequencing consortium of the

CC German Genome Project.

CC s1 sequence also available.

CC This clone (DKFZp434F206) is available at the RZPD in Berlin.

CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,

CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de

XX

FH Key Location/Qualifiers

FH

FT source

1..718

FT

/db_xref="taxon:9606"

FT

/db_xref="RZPD:DKFZp434F206"

FT

/db_xref="UNILIB:1752"

FT

/organism="Homo sapiens"

FT

/clone="DKFZp434F206"

FT

/clone_lib="434 (synonym: htes3). Vector pSport1; host

FT

DH10B; sites NotI + Sali"

FT

/dev_stage="adult"

FT

/tissue_type="testis"

FT

XX

SQ

Sequence 718 BP; 183 A; 196 C; 198 G; 141 T; 0 other;

| | | | | | | |
|-------------|------------|------------|------------|------------|-------------|-----|
| gcttcatgac | agatgttaac | agctggctgc | tcacctttgg | attccagcta | cacaacgtga | 60 |
| tccctgggta | tcccaaacca | gacatggatg | ccatggaacc | ctcctacgag | cttatccaca | 120 |
| cacagatgaa | aacgcaggag | tgggacaaca | gcaagtctat | cctcggggta | cagtgtgaag | 180 |
| tacagaagca | gctcaaggcc | ttgtcacct | tagaacggtt | tgaccagctc | tatggctcca | 240 |
| caatcaccag | ctgccagcag | gctccaaaga | ccaagaagtt | tgcatccagc | ggctcagtct | 300 |
| ttggcaaggg | ggtcaagttt | gccttgaagg | atggccgagt | gaccacagac | atcatcagtg | 360 |
| tggccaatga | ggatgggcga | agggttgctg | ccatcttgaa | ccatgcccac | tacctagaga | 420 |
| acctgcactt | caccattgat | gggttgata | cccattactt | tgtgaaacca | ggaccttcag | 480 |
| aagggtgacct | ggccatcctg | ggcctcagtg | ggggcgcgcg | aaccctggag | aatgggggtca | 540 |
| acgtcactgt | gtcccagatc | aacacagtac | ttaatggcag | gactagacgc | tacacagaca | 600 |
| tccagtccca | gtacggggca | ctgtgcttga | acacacgcta | cgggacaacg | ttggatgagg | 660 |
| agaaggcacg | ggtcctggag | ctgtcccggc | agagagccgt | gcgccaagcg | tgggcccc | 718 |

